

エッジコンピューティング向け AI アクセラレーターエンジンを開発

小型・低消費電力で幅広いアプリケーションを実現

【横浜発、2018年5月11日】 株式会社ソシオネクスト (Socionext Inc.) は、エッジコンピューティング向け AI アクセラレーターエンジン「NNA (Neural Network Accelerator)」を開発しました。本エンジンはディープラーニング処理を高速かつ低消費電力で実現する推論処理の専用アクセラレーターで、画像認識などを行うコンピューター・ビジョン処理において従来型のプロセッサの約 100 倍の性能を発揮します。2018 年第 3 四半期に本 NNA を FPGA に搭載したソフトウェア開発キット (SDK) の提供を開始する予定です。さらに同エンジンを実装した SoC 製品の開発も計画しています。

ソシオネクストは、技術標準化団体クロノス・グループ (Khronos Group) が策定したコンピューター・ビジョン向け API「OpenVX」に対応する VPU (Vision Processor Unit) を搭載したグラフィックス処理用 SoC「SC1810」を提供しています。今回開発した NNA は、この VPU の機能を拡張するアクセラレーターとして開発したもので、車載向けやデジタルサイネージを始めとするさまざまなアプリケーションにおいて、従来型の画像認識処理に加えてディープラーニングを利用したコンピューター・ビジョン処理を高速かつ低消費電力で実現します。

本 NNA にはディープラーニングを実行する上で必要なパラメータやアクティベーションを低ビット化する量子化技術を使った当社独自のアーキテクチャを実装しています。この量子化技術により、少ない演算資源で大量の計算処理が実行でき、かつデータ量を大幅に削減するため、システムメモリバンド幅を抑えることができます。さらに、新しく開発した高効率なデータ供給を行うオンチップメモリ回路技術により、ディープラーニングに重要な演算資源を高い稼働率で動作させることで、非常に小型ながら高いパフォーマンスを達成します。これらの技術により、この NNA を搭載した VPU は画像認識において従来型の VPU の約 100 倍の処理速度を実現できる見込みです。

本 NNA はまず、FPGA に搭載したソフトウェア開発キット (SDK) としての提供を 2018 年第 3 四半期に開始する計画です。また、学習環境として TensorFlow に対応し、量子化技術に対応した専用ライブラリや学習済みモデルから推論処理へのデータ変換ツールを提供します。これら NNA に最適化した学習環境を用いることで、ディープラーニングの低ビット化に必要なモデル圧縮・学習チューニングなどの専門的な知識が無くともモデル構築を効率的に行うことができます。今後はさまざまなディープラーニングフレームワークをサポート

【報道関係者お問い合わせ先】

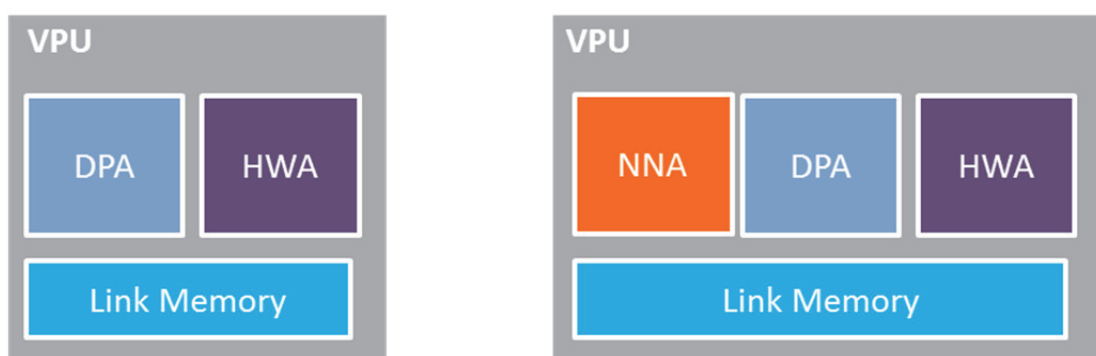
株式会社ソシオネクスト 経営企画室

電話: 045-568-1006

お問い合わせフォーム: <http://socionext.com/jp/contact/>

することで、幅広いアプリケーションに対応できる開発環境を提供する予定です。これにより、ユーザーは簡単にディープラーニングを使ったアプリケーションを構築可能になります。

当社はまたこの NNA を実装した SoC 商品の市場への投入も計画しています。車載システムではカメラ映像から歩行者や自転車などの高精度な物体検知による運転および自動駐車支援、デジタルテレビやサイネージなどのディスプレイシステムでは物体認識によるオブジェクトごとの超解像化や 4K、8K といった高精細化など、その他さまざまなエッジコンピューティング環境で幅広い AI アプリケーションを高性能かつ低消費電力で実現する製品の開発を進めます。



DPA: Data Parallel Accelerator

HWA: Hardware Accelerator

図：従来型(左) および NNA 搭載 (右) VPU のブロック図

製品に関するお問い合わせ

株式会社ソシオネクスト

ビジュアルソリューション事業部

電話：075-778-8011

お問い合わせフォーム：www.socionext.com/jp/contact

ソシオネクストについて

株式会社ソシオネクスト (Socionext Inc.) は、SoC (System-on-Chip) の設計・開発および販売を事業とする新しい企業です。映像・イメージング、ネットワークおよびコンピューティング分野における世界トップレベルの技術を核に、今日のさまざまなアプリケーションの進化を支えます。長年培った技術力と経験、さらに豊富な IP ラインナップをベースに卓越したソリューションを提供し、人々の豊かな体験 = better quality of experience の実現に貢献します。2015年に設立された株式会社ソシオネクストは横浜市に本社を置き、日本国内、アジア、米国およびヨーロッパの各拠点において製品開発および販売活動をグローバルに展開しています。詳しくは socionext.com/jp をご覧ください。

Khronos, OpenVX は Khronos Group Inc.の商標です。

TensorFlow, the TensorFlow logo and any related marks are trademarks of Google Inc.

その他記載されている製品名などの固有名詞は、各社の商標または登録商標です。

プレスリリースに記載された内容、お問い合わせ先などは、発表日現在のものです。その後予告なしに変更されることがあります。あらかじめご了承ください。